The prognostic value of temporal in vitro and in vivo derived hypoxia gene-expression signatures in breast cancer

Radiotherapy and Oncology 102 (2012) 436–443

Contents lists available at SciVerse ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Hypoxia

The prognostic value of temporal *in vitro* and *in vivo* derived hypoxia gene-expression signatures in breast cancer

Maud H.W. Starmans [a,b,*], Kenneth C. Chu [b,c], Syed Haider [b], Francis Nguyen [b], Renaud Seigneuric [a], Michael G. Magagnin [a], Marianne Koritzinsky [a,c,d], Arek Kasprzyk [b], Paul C. Boutros [b], Bradly G. Wouters [a,c,d,1], Philippe Lambin [a,1]

[a] *Department of Radiation Oncology (MaastRO), GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands;* [b] *Informatics and Biocomputing Platform, Ontario Institute for Cancer Research, Toronto, Canada;* [c] *Ontario Cancer Institute and Campbell Family Institute for Cancer Research, Princess Margaret Hospital, University Health Network, Toronto, Canada;* [d] *Departments of Radiation Oncology and Medical Biophysics, University of Toronto, Canada*

ARTICLE INFO

Article history:
Received 3 June 2011
Received in revised form 31 January 2012
Accepted 4 February 2012
Available online 20 February 2012

Keywords:
Gene-expression profiling
Hypoxia
Microarray
Prognostic marker
Princess Margaret Hospital

ABSTRACT

Background and purpose: Recent data suggest that *in vitro* and *in vivo* derived hypoxia gene-expression signatures have prognostic power in breast and possibly other cancers. However, both tumour hypoxia and the biological adaptation to this stress are highly dynamic. Assessment of time-dependent gene-expression changes in response to hypoxia may thus provide additional biological insights and assist in predicting the impact of hypoxia on patient prognosis.

Materials and methods: Transcriptome profiling was performed for three cell lines derived from diverse tumour-types after hypoxic exposure at eight time-points, which include a normoxic time-point. Time-dependent sets of co-regulated genes were identified from these data. Subsequently, gene ontology (GO) and pathway analyses were performed. The prognostic power of these novel signatures was assessed in parallel with previous *in vitro* and *in vivo* derived hypoxia signatures in a large breast cancer microarray meta-dataset ($n = 2312$).

Results: We identified seven recurrent temporal and two general hypoxia signatures. GO and pathway analyses revealed regulation of both common and unique underlying biological processes within these signatures. None of the new or previously published *in vitro* signatures consisting of hypoxia-induced genes were prognostic in the large breast cancer dataset. In contrast, signatures of repressed genes, as well as the *in vivo* derived signatures of hypoxia-induced genes showed clear prognostic power.

Conclusions: Only a subset of hypoxia-induced genes *in vitro* demonstrates prognostic value when evaluated in a large clinical dataset. Despite clear evidence of temporal patterns of gene-expression *in vitro*, the subset of prognostic hypoxia regulated genes cannot be identified based on temporal pattern alone. *In vivo* derived signatures appear to identify the prognostic hypoxia induced genes. The prognostic value of hypoxia-repressed genes is likely a surrogate for the known importance of proliferation in breast cancer outcome.

Radiotherapy and Oncology 102 (2012) 436–443

Poor oxygenation (hypoxia) is a ubiquitous feature of human tumours and exhibits a highly dynamic spatial and temporal distribution both within and between patients [1,2]. Hypoxia is strongly associated with tumour development, growth, invasiveness, resistance to therapy, and contributes to poor prognosis [3–6]. The cellular response to hypoxia is complex and involves several distinct oxygen sensing pathways [1], which have unique activation kinetics and sensitivities to oxygen concentration. Variations in the duration and severity of hypoxic stress therefore lead to substantial phenotypic variations amongst genetically identical cells.

Gene-expression microarray technology provides the possibility to study the function of the whole transcriptome by monitoring the expression of tens of thousands of genes [7]. In oncology, this technique is often used to develop gene-expression profiles from patient tumour tissues as potential prognostic markers. These profiles (gene signatures) can be used to identify different disease subgroups [8] or to predict patient outcome [9,10] or treatment response [11]. Gene-expression signatures might also point to pathways that are important for disease initiation and progression, but extracting the underlying biological mechanisms highlighted by these data remains challenging [7,12]. An alternative approach uses gene-expression profiling of model systems to create

* Corresponding author. Address: Maastricht Radiation Oncology (MaastRO) Lab, University of Maastricht, UNS 50/23 P.O. Box 616, 6200MD Maastricht, The Netherlands.
E-mail address: maud.starmans@maastrichtuniversity.nl (M.H.W. Starmans).
[1] Both authors equally contributed to the study.

0167-8140
doi:10.1016/j.radonc.2012.02.002

experimentally-derived gene signatures that reflect a specific biological process, which can be evaluated to understand the underlying cellular pathways involved [7,10,13,14]. Several of these hypothesis-driven signatures have shown prognostic value [10,13,14].

Several groups, including our own, have identified hypoxia-related gene sets from both *in vitro* [13,15] and *in vivo* data [16,17] that demonstrate prognostic power. The *in vitro* hypoxia signatures tested for prognostic value have been derived so far from a single dataset [13] in which a series of cell lines was exposed to two different hypoxia conditions (0% and 2%) for different periods of time. Originally, a general hypoxia signature derived from this series was shown to have prognostic value [13]. We later showed that acutely responding genes, i.e. those induced at early time-points, had better prognostic value [15]. However, in both cases the performance of these *in vitro* derived signatures was evaluated in a small number of patient gene-expression microarray datasets, studying a limited number of patients. In addition, studies have identified hypoxia-regulated gene sets from other *in vitro* microarray data, however these gene sets were not evaluated in patient data to assess their prognostic value [18,19]. The *in vivo* derived hypoxia signatures were extracted starting from a single known hypoxia up-regulated gene [16] or a set of known hypoxia up-regulated genes [17,20]. Signatures were then created by selecting genes co-regulated with these known hypoxia genes in gene-expression patient data. This approach does not assume that the gene sets identified in this way are directly regulated by hypoxia, but rather that their expression correlates with other genes which are regulated by hypoxia. The *in vivo* derived signatures have also shown high prognostic value on a limited series of patient gene-expression data.

Here, we have investigated the prognostic value of *in vitro* and *in vivo* derived hypoxia gene-expression data in a more comprehensive manner. Previous *in vitro* and *in vivo* hypoxia signatures were tested in a much larger cohort of patients (n = 2311). In addition, we performed a large and comprehensive series of new *in vitro* hypoxia time-course experiments at eight different time-points in three cell lines. We hypothesize that the differences between slow and fast kinetics of hypoxic responses extracted from these data may have differential contributions to patient prognosis. A set of time-dependent expression profiles was used to extract gene-sets that showed different temporal responses to hypoxia. These gene-sets were further studied for their molecular function, involvement in known pathways and their prognostic value in patient datasets and compared to previously defined hypoxia-related signatures. Our data indicate that hypoxia regulated genes identified *in vivo* are superior predictors of patient response.

Material and methods

Cell culture and hypoxic conditions

Exponentially growing prostate (DU145), colon (HT29) and breast (MCF7) carcinoma cells were seeded on glass dishes in McCoy, DMEM or RPMI media, respectively with 10% FCS and transferred to a hypoxic chamber (MACS VA500, Don Whitley Scientific, West-Yorkshire, UK) with atmosphere composition of 5% H_2, 5% CO_2, 0.0% O_2 and residual N_2. Cells were exposed to normoxia (time-point 0 h) and different times of severe hypoxia (1, 2, 4, 8, 12, 16 and 24 h, respectively).

RNA isolation and microarrays

Total RNA was isolated for each time-point using RNeasy (Qiagen, Hilden, Germany) following manufacturer's specifications. RNA quantity/quality was determined using an ND-1000 spectrometer (NanoDrop Technologies, Wilmington, DE) and RNA Nano LabChip kit on the 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA). A summary of these data is provided in Supplementary Tables S1–S3. For each cell line, equal RNA amounts from three independent experiments were pooled for each time-point to minimize experimental variability. Twenty nanograms pooled RNA was processed according to the manufacturer's protocols (Affymetrix, Santa Clara, CA) for 2-cycle amplification and samples were hybridized to Affymetrix HG-U133 Plus 2.0 GeneChips.

Data processing

All analyses were performed in R (v2.12.2). Quality assurance was performed for each time-series. Overall quality is within the manufacturer's standards (Supplementary Tables S4–S6). Data processing was performed for each cell line separately using RMA (affy package, v1.28.0) for pre-processing [21] and updated Entrez GeneID annotation (hgu133plus2hsentrezgcdf package, v14.0.0) [22]. These data are available on www.cancerdata.org and have been deposited in NCBI's Gene Expression Omnibus [23] and are accessible through GEO Series accession number GSE 29641.

An expression filter was applied to remove experimental noise. A low-intensity threshold was set based on the chromosome-Y gene intensities of the two female-derived cell lines (HT29 and MCF7), as described previously [24]. The genefilter package (v1.32.0) was used to remove probes with a $\log_2$-transformed expression value below 5 in at least 6/8 arrays. Only genes passing this filter in all cell lines were retained.

Principal Component Analysis (PCA)

To address whether the cell line datasets should be treated separately or could be analysed as one, a PCA was performed. PCA is tool to reduce the dimensionality of the data. The first three principal components were used to visualize the data and the percentage variance explained was calculated.

Extraction of dynamic gene lists

To extract dynamic gene-expression signatures a multi-step extraction process was performed separately for each cell line. First, genes showing differential expression across the time series were identified. Genes were selected that showed a minimal $\log_2$-transformed fold-change of $(-)1$ (i.e. 2-fold in normal space), in at least two consecutive time-points relative to normoxia. Venn diagrams were drawn to compare cell lines (VennDiagram package, v1.0.0) [25]. Divisive hierarchical clustering was performed using average Euclidean distance linkage (DIANA) (cluster package, v1.13.3). Next, based on the observed expression patterns, seven time-profiles were defined (Table 1). These predefined time-profiles were used as cluster centres in *K*-means clustering (amap package, v0.8-5) to assign each gene to one profile. Finally, the three cell lines were compared to identify genes showing the same time-dependent pattern in two out of three cell lines. Additionally,

Table 1
Different time-profiles used to create gene clusters in relative $\log_2$-transformed units.

	Time-point (h)							
Gene Cluster	0	1	2	4	8	12	16	24
Cluster 1	0	1	1	1	1	1	1	1
Cluster 2	0	0	0	1	1	1	1	1
Cluster 3	0	0	0	0	0	1	1	1
Cluster 4	0	0	0	0	0	−1	−1	−1
Cluster 5	0	0	0	−1	−1	−1	−1	−1
Cluster 6	0	0	0	−1	−1	−1	0	0
Cluster 7	0	−1	−1	−1	−1	−1	−1	−1

two general hypoxia responsive gene lists were generated: genes showing an up-regulation in 2/3 cell lines and genes showing down-regulation in 2/3 cell lines.

Gene ontology (GO)

A GO-analysis was performed for the different gene lists using the GOMiner tool [26]. Standard input settings were used with 1000 randomizations and 5 as smallest category size. Both *p*-values and false discovery rate (FDR) were used to identify over-represented GO-categories. A false discovery rate (FDR) threshold of 0.1 was used to select GO-terms that were enriched per gene cluster.

Pathway analysis

In addition to the GO analysis pathway analyses were performed for each of the different gene lists using the PANTHER software tools [27]. *P*-values were used to identify over- and under-represented pathways.

Testing for prognostic value

Several public mRNA-abundance datasets of primary breast cancer [9,11,28–37] were used to assess whether the dynamic gene lists have prognostic capacity. Dataset pre-processing and annotation were as described for the cell line data, with the addition of housekeeping normalization (to the geometric mean of ACTB, BAT1, B2M and TBP) and median scaling [38]. A multi-gene signature score was then calculated for each gene list as:

$$\text{Score} = \sum_{n=1}^{N} \text{gene}_{\text{expr,n}} \quad (1)$$

Here, N is the number of genes in the list, the parameter $\text{gene}_{\text{expr,n}}$ for a sample is 1 if the sample has a level of gene n above the median for all samples in the dataset and -1 otherwise. This score was used to median dichotomize the patients by dataset and high/low expression patients were pooled across studies. Differences in survival properties between the two groups were assessed with Kaplan–Meier survival curves and Cox proportional hazard ratio modelling followed by the Wald test (survival package, v2.36-5).

To complete the analysis a series of existing hypoxia gene-expression signatures, including the 0% and 2% early signatures [15], 'Chi signature' [13], Elvidge up and down gene sets [18], Sorensen gene set [19], 'Hu signature' [16], the Winter metagene [17] and the Buffa metagene [20], were also tested in the large breast cancer dataset. Methodology was similar to the evaluation of the gene lists as described in the above section.

Results

Hypoxia time series

In order to develop more comprehensive *in vitro* hypoxia gene-expression profiles, we exposed prostate (DU145), colon (HT29) and breast (MCF7) cancer cell lines to severe hypoxia for 0, 1, 2, 4, 8, 12, 16, and 24 h and determined the transcriptome by microarray analysis. Expression of chromosome-Y genes across the three cell lines was used to identify non-expressed genes that were subsequently removed: two cell lines (HT29 and MCF7) are derived from female patients, meaning that intensities of chromosome-Y genes fall within the background noise. Fig. 1A shows that expression of the chromosome-Y genes in the female derived cell lines are, as expected, amongst the lowest intensities, whereas for (male) DU145 the chromosome-Y genes follow the same distribution pattern as the non-chromosome-Y genes.

A PCA was performed to assess the relationship between the transcriptomes in all samples. Fig. 1B displays the first three identified principal components, which encompass 88% of the variability in the dataset. As expected, the three cell lines cluster as individual sets. This conclusion was further supported by a *K*-means ($k = 3$) clustering of samples, which divided them into three groups according to cell line (data not shown). More importantly, the PCA analysis clearly demonstrated the time-series pattern in each of the three datasets with consecutive time-points clustering closer together. Subsequent data analyses to determine hypoxia regulated genes were thus performed separately on each cell line.

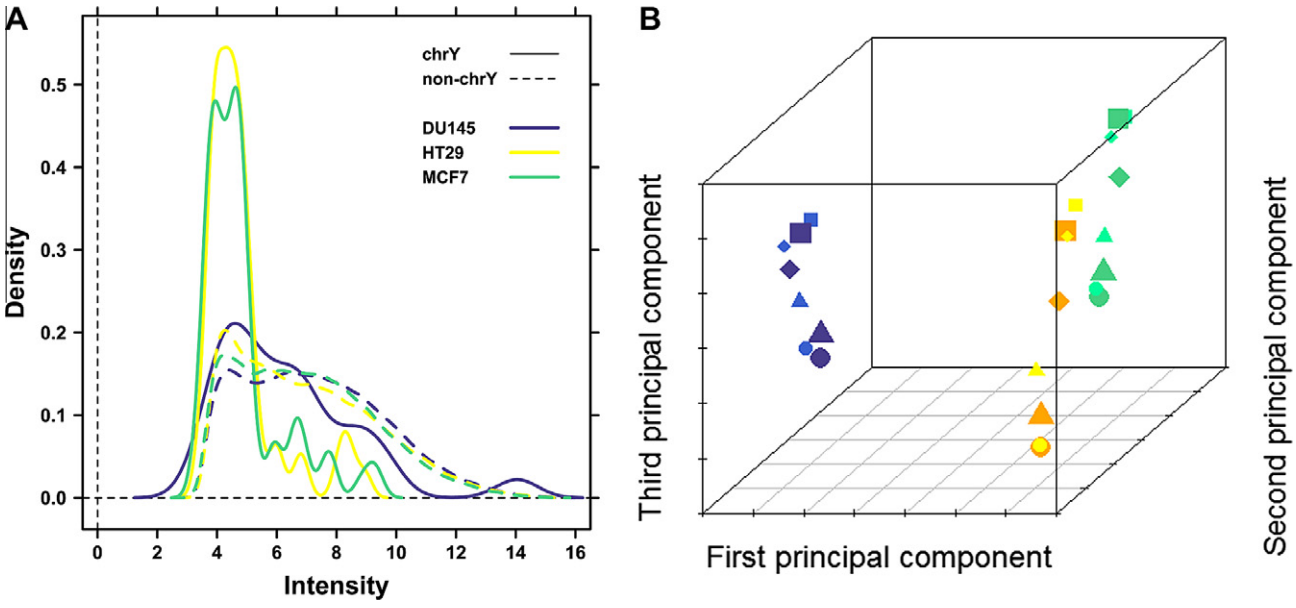

Fig. 1. (A) Gene-expression microarray intensity distributions of chromosome-Y genes (solid curves) and non-chromosome-Y genes (dotted curves) across the time-series in the three cell lines. (B) Scatterplot of the first three principal components from Principal Component Analysis. Each cell line and time-point is represented with a different colour/marker (● 0, • 1, ▲ 2, ▲ 4, ▲ 8, ♦ 12, ■ 16, and ■ 24 h).

Derivation of new hypoxia signatures

In order to identify new hypoxia signatures, we first extracted the genes that showed differential expression across the time series. Each hypoxic time-point was compared to normoxia and genes altered at least 2-fold in two or more consecutive time-points were retained. These genes were clustered to visualize temporal profiles (Supplementary Fig. S1A) and the numbers of selected genes per cell line were compared (Fig. 2A). In Supplementary Fig. S1B and C heatmaps of the three time series are shown for the union and intersection of the selected genes per cell line. Globally, the responses are similar across the three cell lines, however the specific timing and magnitudes of changes are cell line dependent. Furthermore, for subsets of genes a clear response is observed in one/two cell lines, without regulation in the other(s). However, expression changes of the genes selected in all three cell lines are highly similar.

The heatmaps in Supplementary Fig. S1 show that in addition to both induced and repressed genes, there are demonstrable temporal profiles of hypoxia-regulated genes. We defined seven distinct time-profiles (Table 1) and used *k*-means clustering to assign each differentially regulated gene to one of these profiles. Fig. 2B shows heatmaps of the differentially expressed genes per cell lines ordered by cluster. The final seven lists are composed of genes that fall within the same time cluster in at least two cell lines. In addition two global gene lists of induced and repressed genes were created using the same criteria. Table 2 lists an overview of the number of genes in each cluster for each cell line and the number of genes in the final signatures. The content of all nine of these new *in vitro* derived hypoxia signatures are provided in Supplementary Table S7. Gene-expression profiles for seven individual genes, one from each cluster, and average expression profiles per cluster are displayed in Fig. 3 and Supplementary Fig. S2, respectively.

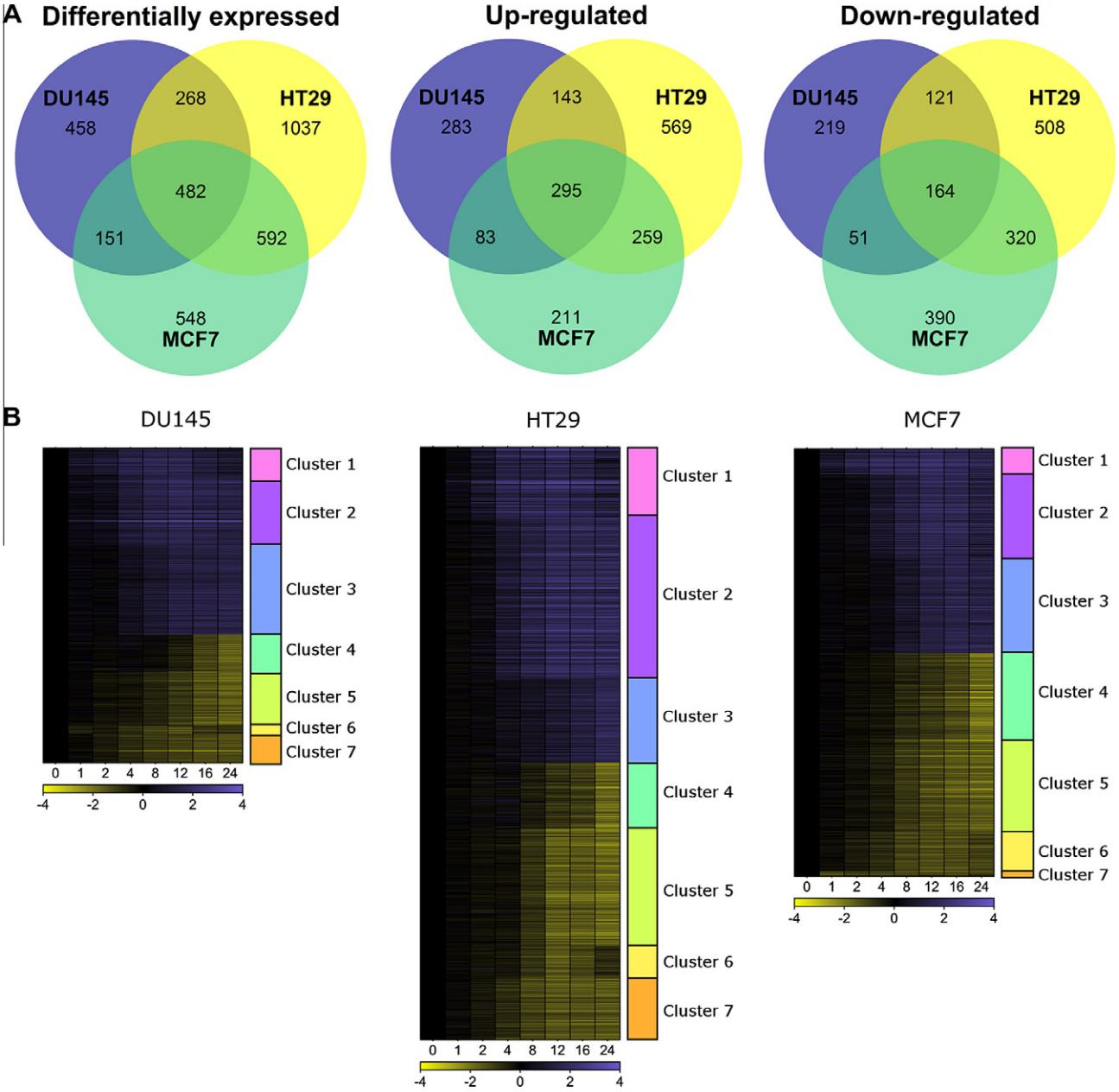

Fig. 2. Identification of time-dependent gene-expression profiles upon hypoxic exposure. Both cell line dependent and independent genes were identified (A). Heatmaps ordered by time-dependent clusters show distinct temporal regulation (B), clusters are shown in the right colour bars.

Table 2
Number of genes identified in each gene profile per cell line and in the final gene lists.

Cell line					
Gene list	DU145	HT29	MCF7	Final list	Fraction overlap: number of genes in final list, divided by the average of the number of genes per cell line
Cluster 1	112	281	107	69	0.41
Cluster 2	306	650	344	246	0.57
Cluster 3	386	335	397	157	0.42
Cluster 4	162	265	356	95	0.36
Cluster 5	233	471	383	162	0.45
Cluster 6	35	129	165	14	0.13
Cluster 7	125	248	21	28	0.21
Up-regulated	804	1266	848	780	0.80
Down-regulated	555	1113	925	656	0.76

Gene ontology (GO) and pathway analysis of new signatures

To provide some biological context to the new signatures we performed GO-analyses for each gene list and the numbers of enriched terms are listed in Supplementary Table S8. A complete list of enriched GO terms is also listed in Supplementary Table S9. GO-terms associated with the induced signatures and global list include apoptosis (GO:0006915, cluster 2, global up) and protein folding (GO:0006457, cluster 3, global up). The down-regulated gene lists are enriched for GO-categories including cell cycle (GO:0007049, cluster 4, 5, global down) and transcription (GO:0006350, cluster 5). A series of GO terms were found in several clusters and in both induced and repressed lists; most of these comprise cellular compartment terms. In addition a significant number of GO-terms were identified with the global induced lists compared to the cluster lists. However many of these terms are either from the cellular compartment category or related to terms identified in the cluster lists. For example the term unfolded protein binding (GO:0051082) is related to the terms response to unfolded protein (GO:0006986) and protein folding (GO:0006457) enriched in cluster 2 and 3, respectively. A pathway analysis was also performed for each separate gene list. Results are provided in Supplementary Table S10 and the pathways significantly over- or under-represented in at least 1 list are shown in Supplementary Fig. S3. The pathways found enriched correspond with the GO analysis. Highest enrichment was found in the apoptosis signalling pathway, angiogenesis and glycolysis for the up-regulated gene lists and cell cycle for the down-regulated ones. Interestingly, p53 signalling was enriched in both up- and down-regulated lists.

Prognostic value of hypoxia gene signatures in breast cancer

Finally, we tested the prognostic potential of the new *in vitro* derived hypoxic signatures and compared them with nine previously published *in vitro* and *in vivo* derived signatures. To do so, we evaluated these signatures in a large number of breast cancer gene-expression datasets comprising 2312 patients. This dataset is many times larger than any other dataset used to evaluate hypoxia gene signatures and thus provides a robust test of their prognostic power. For each signature, a score was calculated for each patient (Eq. (1)), patients were dichotomized into high- and low-score groups, and survival differences assessed. In Fig. 4A and Supplementary Fig. S4A the Kaplan–Meier curves for all of the newly-derived hypoxia signatures are displayed. Surprisingly, none of the signatures consisting of hypoxia-induced genes are prognostic. Similarly, our previously published early hypoxia signatures and the original global hypoxia signature that were prognostic in a smaller dataset, and two other *in vitro* derived hypoxia-induced gene sets are not prognostic in this larger dataset (Supplementary Fig. 4B). Interestingly however, the new repressed gene signatures did show some prognostic value, e.g. cluster 5 dichotomizes patients with a hazard ratio (HR) of 1.35 (95% confidence interval [CI]: 1.18–1.56, $p = 2.17 \times 10^{-5}$, Wald test). Patient classification with the four different significant repressed signatures is similar; more than 75% (1760 patients) of the patients were classified identically by four out of five gene-sets (Fig. 4B). Further, genes identified as hypoxia-repressed by Elvidge et al. [18] showed the same prognostic value (HR: 1.28, CI: 1.11–1.47, $p = 5.16 \times 10^{-4}$, Wald test).

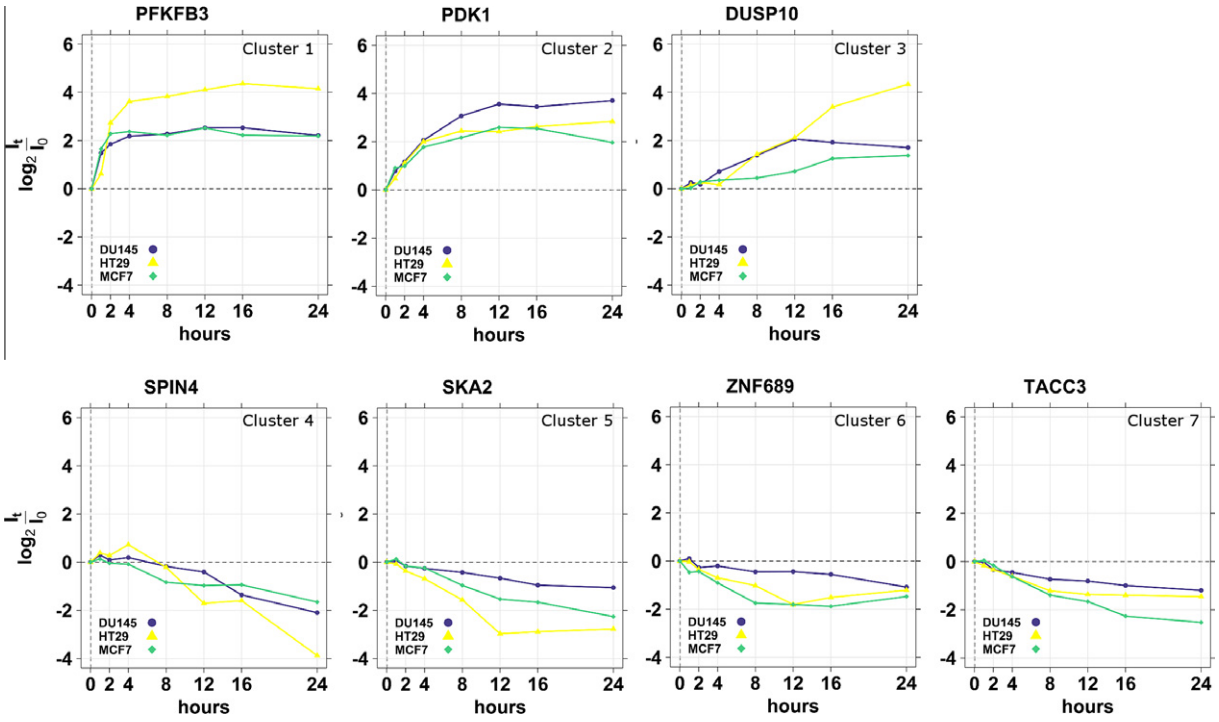

Fig. 3. Gene-expression profiles for seven genes in each of the three cell lines, where each gene is from a distinct cluster.

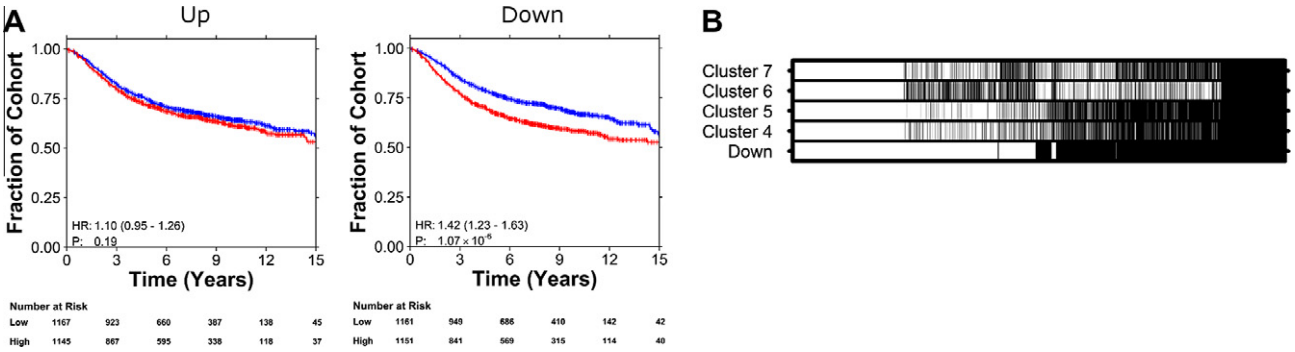

Fig. 4. Evaluation of prognosticity of time-dependent hypoxia responses in a large breast cancer meta-dataset ($n = 2312$). Patients were assigned to low-score (blue curves) and high-score (red curves) groups based on genes in global regulation lists (A) and Kaplan–Meier survival curves were drawn. Hazard ratios [HR] and p-values are from Cox proportional hazard ratio modelling followed by the Wald test. Patient classification with the five new prognostic gene lists shows large agreement (B). White/black indicate patients assigned to the low-score/high-score group, respectively.

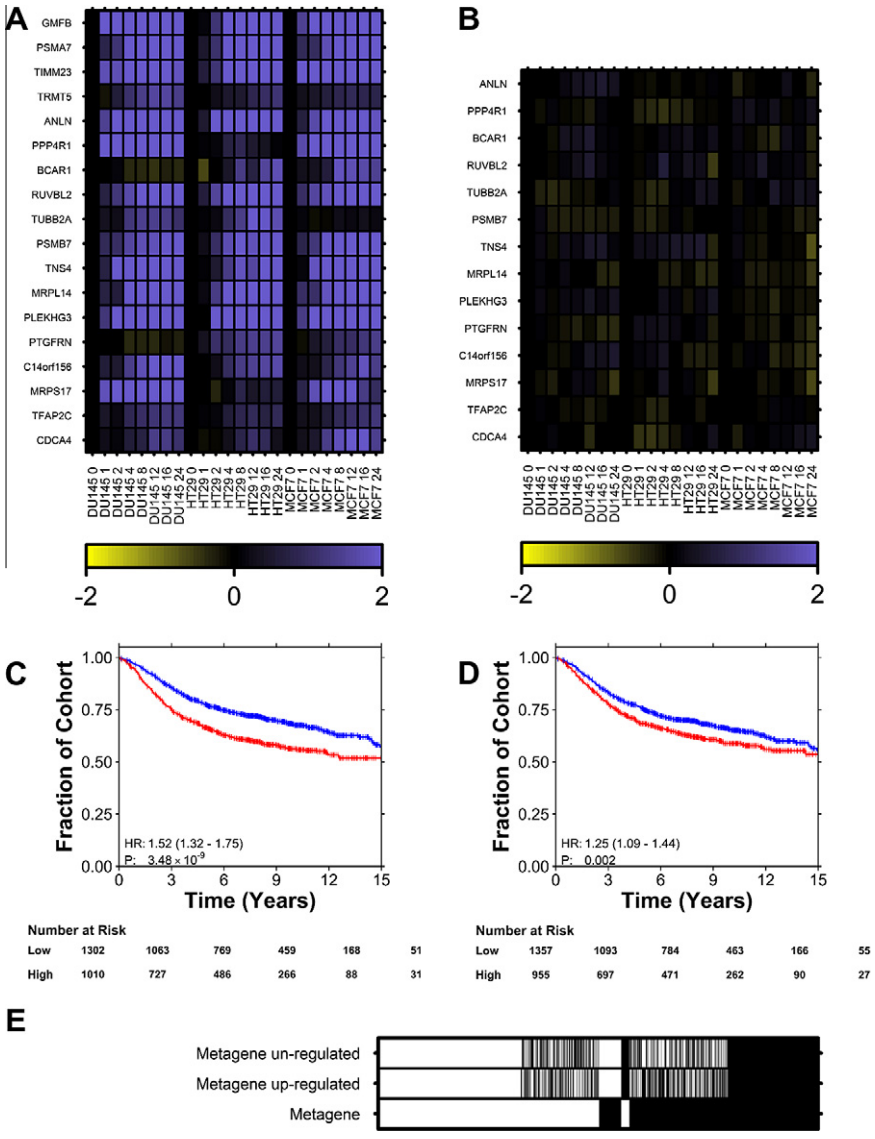

Fig. 5. Expression of subset of the Winter metagene that were selected as up-regulated (A) or as un-regulated (B) by hypoxia *in vitro* and evaluation of prognosticity of these subsets in a large breast cancer meta-dataset ($n = 2312$). Patients were assigned to low-score (blue curves) and high-score (red curves) groups based on genes the up-regulated subset (C) or un-regulated subset (D). Patient classification with the full Winter metagene, up-regulated and un-regulated subset show distinct patterns (E).

We also tested previously published *in vivo* derived hypoxia signatures consisting of genes whose expression *in vivo* correlates with known hypoxia-induced genes. In sharp contrast to the *in vitro* derived signatures, these signatures show high prognostic

power in the large breast cancer meta-dataset (Supplementary Fig. 4C). The first published signature (Winter metagene), demonstrated a hazard ratio of 1.58 ($p = 1.38 \times 10^{-10}$, Wald test). To investigate this discrepancy further, we evaluated the expression of genes contained in the *in vivo* signatures in our new dataset. For all three *in vivo* signatures, a large fraction of the genes show clear up-regulation in the time-series (Supplementary Fig. S5). However, a number of genes, especially in the two larger gene lists, are not induced by hypoxia in any of the cell lines and a subset is even consistently repressed. We therefore analysed separately the prognostic power of genes that showed clear evidence of hypoxic regulation *in vitro* from those that did not of one of these signatures (Fig. 5A and B). Importantly, we find that the hypoxia-regulated genes in the *in vivo* signature retain clear prognostic power, even when removed from those that were not regulated by our criteria within the new dataset (Fig. 5C). However, the subset of genes in the *in vivo* signature that were not hypoxia-regulated also contribute to the prognostic power of this signature (Fig. 5D). This is also reflected in the patient classification (Fig. 5E), which is substantially different when using either of the subsets or the complete signature.

Discussion

In order to better define the potential prognostic value of gene-expression signatures we have extended previous reported results in three significant ways. First, we conducted the largest and most detailed transcriptome response to hypoxia published to date in order to derive novel *in vitro* hypoxia signatures that better reflect the nature of this stress. This dataset consists of three diverse cell lines analysed at multiple time-points that allow expression of genes regulated by different hypoxia responsive signalling pathways [1]. In both published and unpublished data, we have validated multiple hypoxia-regulated genes identified from these time-series using quantitative RT-PCR [5,39,40]. Globally, responses are similar amongst the lines and for a subset of genes, time-patterns are also similar. Discrepancies amongst the lines with respect to meeting our criteria for regulation largely reflect differences in magnitude of regulation, rather than in regulation direction (Supplementary Fig. 1B). To some extent, the temporal pattern is cell line dependent, as evidenced by a larger overlap between global gene lists than for individual signatures (Table 2). The dynamic natures of these data highlight the limitations of employing a single time-point or cell line for determination of hypoxia-responsive genes. This led us to extract a series of seven different signatures consisting of genes with distinct temporal profiles to evaluate in patient derived gene-expression data. Gene ontology and pathway analyses suggest that each signature is likely to contain genes that respond to both common and unique biological processes including many previously implicated in hypoxia. Terms linked to the unfolded protein response, ER stress, autophagy, cell death and metabolism [1,41,42] were enriched in the induced signatures, whereas terms linked to proliferation, RNA processing and splicing were found in the repressed gene signatures [43].

Second, we have evaluated our new temporal signatures with other previously published *in vitro* and *in vivo* hypoxia signatures in a much larger cohort of breast cancer patients than done previously. Surprisingly, we found that even with the inclusion of the new signatures, none of the *in vitro* derived signatures consisting of hypoxia-induced genes are prognostic. This result contrasts with previous studies [13,15–17], including our own focusing on early responding hypoxia induced genes. This result stresses the importance of validation, and likely results from the fact that the number of patients evaluated in these early studies was limited to a few hundred. In this study we combined a large number of datasets to create a highly powered dataset, comprising data on more than 2300 breast cancer patients.

The failure of these *in vitro* signatures contrasts with the strong prognostic performance of the three *in vivo* derived hypoxia signatures, which were clearly able to split patients in groups with different survival properties (Supplementary Fig. 4). This was somewhat surprising given that the *in vivo* derived signatures are derived from and consist of several known hypoxia-induced genes [16,17,20]. One possibility was that the *in vivo* signatures reflected biological characteristics distinct from hypoxia itself. For example, the *in vitro* studies do not capture the more complex tumour microenvironment associated with hypoxia. Tumour hypoxia is often associated with changes in pH, nutrient gradients, and waste gradients that can influence gene-expression in important ways [19,44–46]. It is also possible that tumour hypoxia *in vivo* is associated with over-expression of specific genes that themselves are not hypoxia regulated. Such genes could, for example, influence the tolerance of hypoxic cells. However, it is unlikely that this is the primary basis for the difference in the prognostic value of the signatures. Our analysis demonstrates clearly that the hypoxia-regulated genes in the *in vivo* signature contribute to the prognostic power of the signature (Fig. 5C). Since these hypoxia-regulated genes are included within several of our new signatures, we must conclude that many other of the co-regulated hypoxia-induced genes in the *in vitro* signatures do not contribute to the prognostic value of hypoxia. Thus, it will be important to identify the fraction of 'clinically relevant' hypoxia regulated genes from the larger fraction of co-regulated genes *in vitro* for use as a predictor of patient prognosis.

Finally, our results also indicate that signatures derived from *in vitro* hypoxia repressed genes also exhibit prognostic power. Hypoxia repressed genes and their clinical importance have not been assessed to the same extent as hypoxia-induced genes. This is largely due to the perceived importance of the HIF pathway in hypoxia, which influences cell biology primarily through increased transcription [47,48]. However, hypoxia is also an important suppressor of cellular metabolism through mTOR and UPR pathways [49–51] and this is often associated with repressed gene-expression and reduced proliferation. It is unlikely that the prognostic power of these signatures is due to hypoxia itself. Our data indicate that high expression of the repressed genes in these signatures is associated with worse survival, which is the opposite of the expected influence of tumour hypoxia. The GO and pathway analyses revealed that a large fraction of the genes in these repressed signatures are involved in cell cycle, suggesting that the signatures may rather be surrogate markers for proliferation. This agrees with numerous previous publications implicating proliferation in breast cancer as a major contributor to the prognostic value of gene-expression microarray-based signatures [14,52].

Acknowledgements

This work was financially supported by Siemens (M.H.W.S.), the CTMM framework (AIRFORCE project to M.H.W.S. and P.L.), the Dutch Cancer Society (KWF Grant UM 2008-4068 to B.W.), the Ontario Ministry of Health and Long Term Care (OMOHLTC), the Terry Fox New Frontiers Research Program (PPG09-020005 to B.W. and M.K.), the Ontario Institute for Cancer Research and Terry Fox Research Institute (Selective therapies program to B.W.), the Canadian Institute for Health Research (CIHR Grant 201592 to B.W. and M.K.) and the EU 6th and 7th framework programs (METOXIA Project 222741 to B.W., P.L. and M.K.). The views expressed do not necessarily reflect those of the OMOHLTC. This study was also conducted with the support of the Ontario Institute for Cancer Research (P.C.B. and B.W.) through funding provided by the Government of Ontario.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.radonc.2012.02.002.

References

[1] Wouters BG, Koritzinsky M. Hypoxia signalling through mTOR and the unfolded protein response in cancer. Nat Rev Cancer 2008;8:851–64.
[2] Dewhirst MW, Cao Y, Moeller B. Cycling hypoxia and free radicals regulate angiogenesis and radiotherapy response. Nat Rev Cancer 2008;8:425–37.
[3] Hockel M, Vaupel P. Biological consequences of tumor hypoxia. Semin Oncol 2001;28:36–41.
[4] Brown JM, Wilson WR. Exploiting tumour hypoxia in cancer treatment. Nat Rev Cancer 2004;4:437–47.
[5] Rouschop KM, van den Beucken T, Dubois L, et al. The unfolded protein response protects human tumor cells during hypoxia through regulation of the autophagy genes MAP1LC3B and ATG5. J Clin Invest 2010;120:127–41.
[6] Overgaard J. Hypoxic modification of radiotherapy in squamous cell carcinoma of the head and neck – a systematic review and meta-analysis. Radiother Oncol 2011;100:22–32.
[7] Bild AH, Potti A, Nevins JR. Linking oncogenic pathways with therapeutic opportunities. Nat Rev Cancer 2006;6:735–41.
[8] Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406:747–52.
[9] Miller LD, Smeds J, George J, et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. Proc Natl Acad Sci USA 2005;102:13550–5.
[10] Chang HY, Sneddon JB, Alizadeh AA, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS Biol 2004;2:E7.
[11] Symmans WF, Hatzis C, Sotiriou C, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. J Clin Oncol 2010;28:4111–9.
[12] Quackenbush J. Microarray analysis and tumor classification. N Engl J Med 2006;354:2463–72.
[13] Chi JT, Wang Z, Nuyten DS, et al. Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. PLoS Med 2006;3:e47.
[14] Starmans MH, Krishnapuram B, Steck H, et al. Robust prognostic value of a knowledge-based proliferation signature across large patient microarray studies spanning different cancer types. Br J Cancer 2008;99:1884–90.
[15] Seigneuric R, Starmans MH, Fung G, et al. Impact of supervised gene signatures of early hypoxia on patient survival. Radiother Oncol 2007;83:374–82.
[16] Hu Z, Fan C, Livasy C, et al. A compact VEGF signature associated with distant metastases and poor outcomes. BMC Med 2009;7:9.
[17] Winter SC, Buffa FM, Silva P, et al. Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. Cancer Res 2007;67:3441–9.
[18] Elvidge GP, Glenny L, Appelhoff RJ, Ratcliffe PJ, Ragoussis J, Gleadle JM. Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1alpha, HIF-2alpha, and other pathways. J Biol Chem 2006;281:15215–26.
[19] Sorensen BS, Toustrup K, Horsman MR, Overgaard J, Alsner J. Identifying pH independent hypoxia induced genes in human squamous cell carcinomas in vitro. Acta Oncol 2010;49:895–905.
[20] Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. Br J Cancer 2010;102:428–35.
[21] Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249–64.
[22] Dai M, Wang P, Boyd AD, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res 2005;33:e175.
[23] Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10.
[24] Bachtiary B, Boutros PC, Pintilie M, et al. Gene expression profiling in cervical cancer: an exploration of intratumor heterogeneity. Clin Cancer Res 2006;12:5632–40.
[25] Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics 2011;12:35.
[26] Zeeberg BR, Feng W, Wang G, et al. GoMiner: a resource for biological interpretation of genomic and proteomic data. Genome Biol 2003;4:R28.
[27] Thomas PD, Kejariwal A, Guo N, et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucleic Acids Res 2006;34:W645–50.
[28] Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res 2005;7:R953–64.
[29] Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365:671–9.
[30] Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 2006;439:353–7.
[31] Chin K, DeVries S, Fridlyand J, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 2006;10:529–41.
[32] Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 2006;98:262–72.
[33] Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res 2007;13:3207–14.
[34] Bos PD, Zhang XH, Nadal C, et al. Genes that mediate breast cancer metastasis to the brain. Nature 2009;459:1005–9.
[35] Zhang Y, Sieuwerts AM, McGreevy M, et al. The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. Breast Cancer Res Treat 2009;116:303–9.
[36] Sabatier R, Finetti P, Cervera N, et al. A gene expression signature identifies two prognostic subgroups of basal breast cancer. Breast Cancer Res Treat 2011;126:407–20.
[37] Schmidt M, Petry IB, Bohm D, et al. Ep-CAM RNA expression predicts metastasis-free survival in three cohorts of untreated node-negative breast cancer. Breast Cancer Res Treat 2011;125:637–46.
[38] Barsyte-Lovejoy D, Lau SK, Boutros PC, et al. The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. Cancer Res 2006;66:5330–7.
[39] Mujcic H, Rzymski T, Rouschop KM, et al. Hypoxic activation of the unfolded protein response (UPR) induces expression of the metastasis-associated gene LAMP3. Radiother Oncol 2009;92:450–9.
[40] Ramaekers CH, van den Beucken T, Meng A, et al. Hypoxia disrupts the Fanconi anemia pathway and sensitizes cells to chemotherapy through regulation of UBE2T. Radiother Oncol 2011;101:190–7.
[41] Rouschop KM, Ramaekers CH, Schaaf MB, et al. Autophagy is required during cycling hypoxia to lower production of reactive oxygen species. Radiother Oncol 2009;92:411–6.
[42] van den Beucken T, Ramaekers CH, Rouschop K, Koritzinsky M, Wouters BG. Deficient carbonic anhydrase 9 expression in UPR-impaired cells is associated with reduced survival in an acidic microenvironment. Radiother Oncol 2009;92:437–42.
[43] Koritzinsky M, Rouschop KM, van den Beucken T, et al. Phosphorylation of eIF2alpha is required for mRNA translation inhibition and survival during moderate hypoxia. Radiother Oncol 2007;83:353–61.
[44] Sorensen BS, Alsner J, Overgaard J, Horsman MR. Hypoxia induced expression of endogenous markers in vitro is highly influenced by pH. Radiother Oncol 2007;83:362–6.
[45] Brahimi-Horn MC, Bellot G, Pouyssegur J. Hypoxia and energetic tumour metabolism. Curr Opin Genet Dev 2011;21:67–72.
[46] Voss MJ, Niggemann B, Zanker KS, Entschladen F. Tumour reactions to hypoxia. Curr Mol Med 2010;10:381–6.
[47] Majmundar AJ, Wong WJ, Simon MC. Hypoxia-inducible factors and the response to hypoxic stress. Mol Cell 2010;40:294–309.
[48] Semenza GL. Oxygen homeostasis. Wiley Interdiscip Rev Syst Biol Med 2010;2:336–61.
[49] Guertin DA, Sabatini DM. Defining the role of mTOR in cancer. Cancer Cell 2007;12:9–22.
[50] Koritzinsky M, Magagnin MG, van den Beucken T, et al. Gene expression during acute and prolonged hypoxia is regulated by distinct mechanisms of translational control. EMBO J 2006;25:1114–25.
[51] Koumenis C, Naczki C, Koritzinsky M, et al. Regulation of protein synthesis by hypoxia via activation of the endoplasmic reticulum kinase PERK and phosphorylation of the translation initiation factor eIF2alpha. Mol Cell Biol 2002;22:7405–16.
[52] Wirapati P, Sotiriou C, Kunkel S, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Res 2008;10:R65.